A program for fractal and multifractal analysis of two-dimensional binary images: Computer algorithms versus mathematical theory

Edith Perrier , Ana M. Tarquis , Annette Dathe

Abstract

In this paper we present a tool to carry out the multifractal analysis of binary, two-dimensional images through the calculation of the Rényi $D(q)$ dimensions and associated statistical regressions. The estimation of a (mono)fractal dimension corresponds to the special case where the moment order is $q=0$.

1. Introduction

Fractal and multifractal concepts have been increasingly applied in various fields of science. The purpose of the software developed here is to enable researchers to use a classical multifractal analysis algorithm with their own data, and to be able to overcome some numerical drawbacks by carrying out their own numerical experiments.

2. Mathematical theory and computer algorithm

2.1. Theory

We refer to Dathe et al. (2006-this issue) for an up to date bibliography and for many details about multifractal theory. Here we will recall the classical definition of the Rényi dimensions of order q, which are commonly used to characterize a multifractal object embedded in a d-dimensional space. Let us denote $p_i(r)$ the object density in any box i of size r covering a part of the object:

$$p_i(r) = \frac{M_i(r)}{M} \tag{1}$$

where $M_i(r)$ is the mass of the object included in box i and M is the total mass of the object. The concept of generalized dimensions $D(q)$ corresponds to the scaling exponents for the qth moment of the density. They are defined as:

$$D(q) = \lim_{r\to 0} \frac{1}{q-1} \frac{\log \sum_{i=1}^{N(r)} p_i^q(r)}{\log r} \tag{2}$$

where $N(r)$ is the number of boxes used to cover the object, and for $q=1$ as:

$$D(1) = \lim_{r\to 0} \frac{\sum_{i=1}^{N(r)} p_i(r) \log p_i(r)}{\log r} \tag{3}$$

In practice, the object density as defined in Eq. (1) is taken to the respective power of q, summed for all i, and

plotted versus the box size in a log–log coordinate system. From the slope, which is also called the mass exponent τ, the generalized dimensions are estimated as $D(q)=\tau(q)/(1-q)$.

2.2. Algorithm

For 2-dimensional images, $M_i(r)$ is the number of pixels included in a squared box of linear size r. In the following source code, for each box size r, $M_i(r)$ is denoted by Mr[i] for i varying from 0 to NumberOfBoxes-1 (that is from 1 to $N(r)$).

```
public void CalculateDqs (int q,int minBox,int
maxBox, int RMAX)
{
double x=0, y=0, slope;
int l, c;
for(IndexBox=minBox;IndexBox<maxBox+1;
IndexBox++)
   {
// BOX ACTUAL SIZE
   r=Math.pow(2,(double)(IndexBox));
   int rint=(int)Math.round(r);
// INITIALISATIONS
   int NumberOfBoxesLin=(int)(RMAX/rint); //
SQUARED IMAGE
   int NumberOfBoxes NumberOfBoxesLin*
NumberOfBoxesLin;
   int Mr[]=new int[NumberOfBoxes];
   double p[]=new double [NumberOfBoxes];
   int i;
   for (i=0;i<NumberOfBoxes;i++)
   {Mr[i]=0;p[i]=0.0;}
   //ADJUSTING THE MAXIMUM BOX SIZE
maxR TO THE SIZE RMAX OF THE SELECTED
IMAGE
   int M=0;
   int maxR=RMAX
   if(NumberOfBoxesLin*rint<RMAX) maxR=
NumberOfBoxes*rint;
//COUNTING THE PIXELS IN EACH BOX FROM
THE IMAGE ARRAY
   for(l=0;l<maxR;l++)
   for(c=0;c<maxR;c++)
      {
      int nl=(l-l%rint)/rint;
      int nc=(c-c%rint)/rint;
      if (imagearray[l][c]==0)
      Mr[NumberOfBoxesx*nc+nl]=Mr[Number-
OfBoxesx*nc+nl]+1;
      }
// FROM MASSES Mi TO DENSTIES Pi
   for(i=0; i<NumberOfBoxes;i++) M=M+Mr[i];
   for (i=0;i<NumberOfBoxes;i++) p[i]=Mr[i]/
(double)M;
//CALCULATING EQS.1 and 2 NUMERATORS
   double q_moment=0;
   for (i=0;i<NumberOfBoxes;i++)
   if (p[i]>0)
      {
      if (q==1)
      q_moment=q_moment+(p[i]*Math.log(p[i]));
      else
      q_moment=q_moment+Math.pow(p[i],q);
      }
// PREPARING THE NUMERATORS y VS DE-
NOMINATOR x FITS
   x=Math.log(RMAX/NumberOfBoxesLin);
// when the image size RMAX is a multiple of the box
size r,
// x REDUCES TO x=Math.log(r)
   if (q==1)
   y=q_moment;
   else
   y=Math. log (q_moment);
   PlotPoint Pxy= new PlotPoint (x,y);
   ListPlotPoints.addPlotPoint (Pxy);
   }
//CALCULATING THE LIMIT IN EQS.1 and 2
WHEN TENDS TOWARDS 0 IS REPLACED BY //
ESTIMATIONS OF THE SLOPE OF THE BEST y/
x REGRESSION LINE.
Slope=CalculateRegress(listPlotPoints);
//FINALLY ACCORDING TO EQS.1 and 2
DEFINITIONS
if(q==1)y=slope; else y=slope/(double)(q-1);
Dq=slope;
}
```

2.3. Using the software

One can use the software with any square black-and-white image that is provided in the "jpeg" format. If the image is given with a gray scale ranging from 0 to 255, the values are automatically binarized using a simple threshold value equal to 127. Images can be preprocessed using any image processor, e.g to convert from color/gray images to black-and-white using different types of algorithms, to invert background and foreground, to extract boundaries, etc. The program can be run using any computer. There are two options to run it: either using the applet version (MFapplet.java) (this applet can be also run distantly on the Web) or using the

local application version (MFmain.java). The local application does require installation of java tools (free from http://java.sun.com/) but enables the selection of an input image and the saving of numerical results in text files. All the installation details are given on the webpage http://www.bondy.ird.fr/~perrier/multifractals/MF.htm, where all the codes (including both MFapplet.java and MFmain.java) can also be downloaded.

Fig. 1 describes the program interface which appears on the computer screen as soon as the program has been launched. The "select image" button gives access to the user's file system (available only in the local application version) for selection of any "imageName".jpg image; the "draw image" button displays the image on the screen. The computations can be done separately for every q, thus enabling the user to select the moment order (any q value from −10 to 10) and to look at the successive grids covering the image during computation. The user can select the range of box sizes using the ruler bars, which give the smallest and biggest box size as powers of base 2, 3, 4, or 5. The successive box sizes r are used to calculate the regression (x, y) data (from Eqs. (2) and (3), $x=\log r$, $y=\sum_{i=1}^{n(r)} p_i(r)\log p_i(r)$ if $q=1$, else $y=\log\sum_{i=1}^{n(r)} p_i(r)^q$). Values of $D(q)$ for the whole range of q can also be calculated in a single run and several trial spectra can be superimposed as shown in Fig. 1. The "Save" button closes a text file named imageName.RESULTS.dat (from the user-selected image file name), which includes all the numerical results obtained from the beginning of the simulation (or from the last click on the "Clear" button). A user can examine regression linearity and residuals and see the coefficient of determination R^2 obtained for different number of points and associated *d.d.l.* (degree of freedom) in the regression. Further analysis of the results can be done by importing the text file into any dedicated statistical software.

3. Results with the test imagery

3.1. Six example images

Six examples are included to illustrate the software (Fig. 2).

3.1.1. The Sierpinski Carpet

This example is obtained by dividing a square into nine subsquares, removing the central square, and repeating the same type of division in a self-similar way ad infinitum. The fractal dimension D is $D(0)=(\log 8/\log 3)$. The image shown in Fig. 2a is an approximation of the Sierpinski Carpet obtained after six iterations.

3.1.2. The PSF model

The PSF is also defined (Perrier et al., 1999) by an iterative process: dividing a square into n^2 subsquares; then generating constant proportions s, p and f of solids, pores and porous subparts within these subsquares; then continuing across an infinite range of scales by partitioning the f porous subsquares using the same proportions as for the whole. The fractal dimension D is $D(0)=(\log(n^2 f)/\log n)$. D is the fractal dimension of the interface, i.e. of the boundary between void and solid spaces, which is a fractal curve (if $d=2$) or a fractal surface (if $d=3$). It has been shown that the mass of the solid space and the volume of the pore space are not fractal ($D=d$). The PSF image in Fig. 2c is defined by the parameters $n^2=4^2$, $f=8/16$ and $s=5/16$ (thus $p=3/16$). It is a finite approximation of the PSF model obtained after 4 iterations. The black object representing the solid mass in Fig. 2c is not fractal ($D=2$). Fig. 2b depicts the pore solid interface of the black and white image displayed in Fig. 2c. This is a finite approximation of a fractal object of dimension $D(0)=(\log 8/\log 4)=1.5$.

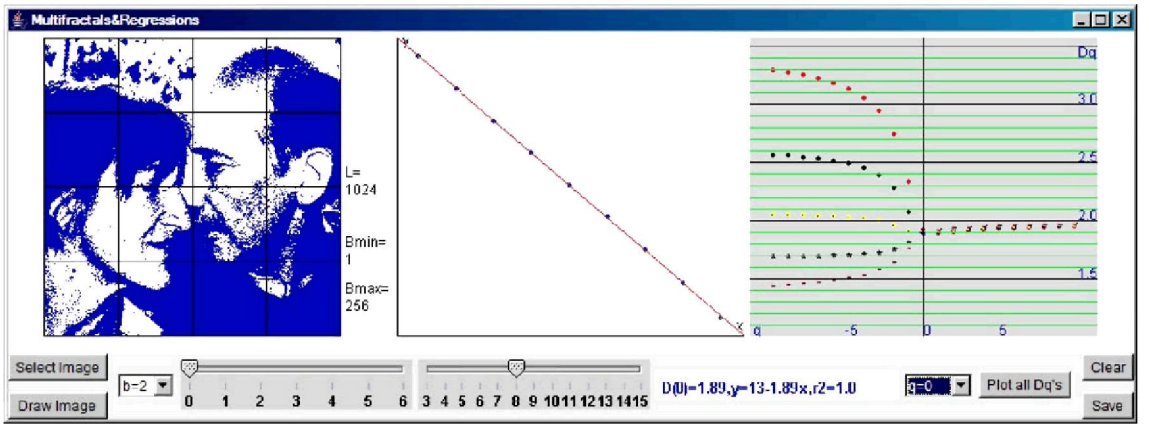

Fig. 1. The software interface. Results on an arbitrary image for several box size ranges.

3.1.3. The multifractal MFp_1p_2 model

The multifractal model depicted by Stanley and Meakin (1988) involves two parameters, p_1 and p_2. Here $p_1=1$ and $p_2=0.5$ (Fig. 2d). A square is divided into 4 squares of linear size 1/4, and the process is repeated with $k=8$ iterations. The set of black pixels is obtained by successive Euclidean divisions, it is non-fractal ($D(0)=2$) but the spatial repartition of the mass is non-homogeneous. This set of black pixels is the result of a spatial multiplicative process where, inside each box of size r, the probability for a pixel to be black to is $p_i(k) = \prod_{j=1}^{k} p_j$ where $p_j=p_1$ or $p_j=p_2$.

3.1.4. Two images of a soil thin section

These images are fully described in Dathe et al. (2006-this issue). They show respectively the pore solid interface (Fig. 2e) and the solid space (Fig. 2f) of a real soil thin section denoted SEM3. The fractal dimensions calculated by Dathe and Thullner (2005) using their own box-counting codes are $D(0)=1.53$ for the SEM3 interface and $D(0)=1.94$ for the SEM3 solid mass.

3.2. Theoretical results

The PSF interface is fractal, but its mass is neither fractal nor multifractal and

$$D(q) = d \tag{4}$$

for any q. We give in Appendix 1 a full mathematical proof of this fact (see Eq. (A1.8)) that replicates the algorithm using a mathematical language, which provides another way to understand the code. We compared the mathematical values to those obtained by the computer algorithm: they are identical when we restrict the mathematical calculations over the same finite box range size selection as those used by the algorithm.

The multifractal image MFp_1p_2 depicted by Stanley and Meakin (1988) can also be analyzed mathematically.

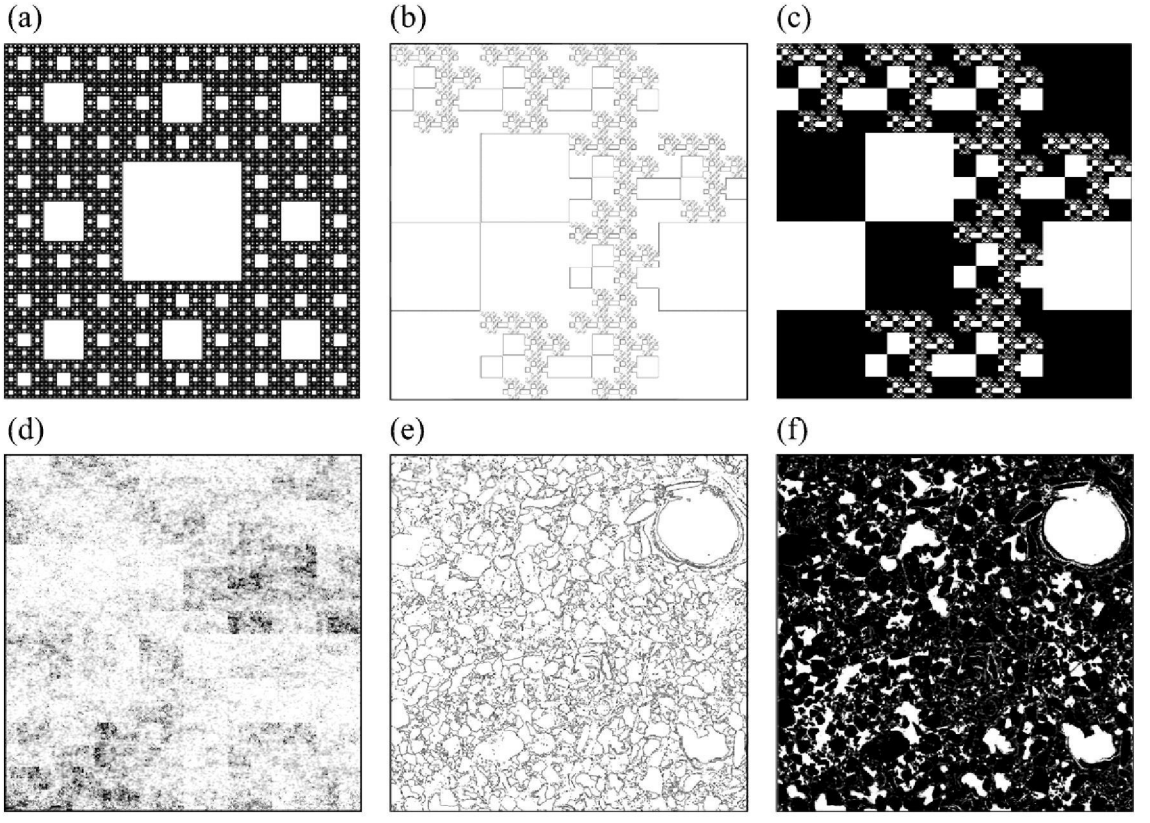

Fig. 2. Six example images. a) Sierpinski Carpet, size: 729 * 729 pixels. Nine generation iterations. b) PSF4Interface, size: 1024 * 1024 pixels. $n^2=4^2$, $p=3/16$, $s=5/16$, ($f=8/16$), 4 iterations. c) PSF4SolidMass, size: 1024 * 1024 pixels. The object is the solid mass colored in black. b shows the interface between this object and its dual (that is the void "mass"). d) An MFp_1p_2 multifractal image, size: 1024 * 1024 pixels. A random realization created with Stanley and Meakin's (1988) model. $p_1=1$, $p_2=0.5$, $k=8$ iterations in a square divided by 4^8. e) SEM3Interface, size: 1024 * 1024 pixels. Void–solid interface from a thin section of Luvisol developed on Loess at the μm scale (see Dathe et al., 2006-this issue). f) SEM3SolidMass, size: 1024 * 1024 pixels. Solid mass colored black, using the same image as in e).

We show in Appendix 2 that the fractal dimensions of this hierarchical geometrical construction are (see Eqs. (A2.3) and (A2.4)):

$$D(q) = 1 + \frac{\log(p_1^q + p_2^q) - q\log(p_1 + p_2)}{(1-q)\log(2)} \quad \text{for } q \neq 1, \tag{5a}$$

$$\text{and } D(1) = \frac{\log(2(p_1 + p_2)) - (p_1 \log p_1 + p_2 \log p_2)}{\log(2)} \tag{5b}$$

Regarding the fractal interface D in the PSF model, or the mass fractal dimension D of the Sierpinski Carpet, the expected theoretical result is

$$D(q) = D \tag{6}$$

for any q.

3.3. Numerical results

Fig. 3 gives the log/log regression fit and the estimated value of $D(0)$ for each image, using the present computer program. We compare them to the theoretical ones or to the ones already measured on real data.

The theoretical value of the fractal dimension for the Sierpinski Carpet is $D = \log 8/\log 3$. With box sizes ranging from 1 to the 729 pixel image size, and using as successive box sizes powers of the 3 ratio (Fig. 3a), as used in the theoretical model, the value of $D(0)$ is estimated by the numerical algorithm is 1.8927892607143721. This is an exact result with sixteen digits.

The theoretical value for the PSF model approximated by the four level PSF4Interface is $D(0) = \log 8/\log 4 = 1.5$. Using boxes from 2 to 128 pixel linear size to measure the interface object gives an estimated value of 1.5 (Fig. 3b), with a coefficient of determination R^2 equal to 0.996 for the 7 points (that is 6 *d.d.l.*) in the log/log linear regression fit.

The theoretical mass "fractal" dimension of the PSF solid mass is $D = 2$. The estimated value for the PSF4SolidMass image is $D(0) = 1.95$, the log/log linear fit appears almost perfect (Fig. 3c) and the associated coefficient of determination R^2 equals 0.99988. The reason for this discrepancy is the finite size of the image. Perrier et al. (1999) had already indicated that some images may appear to be mass fractal even if we know that they are not actual mass fractals.

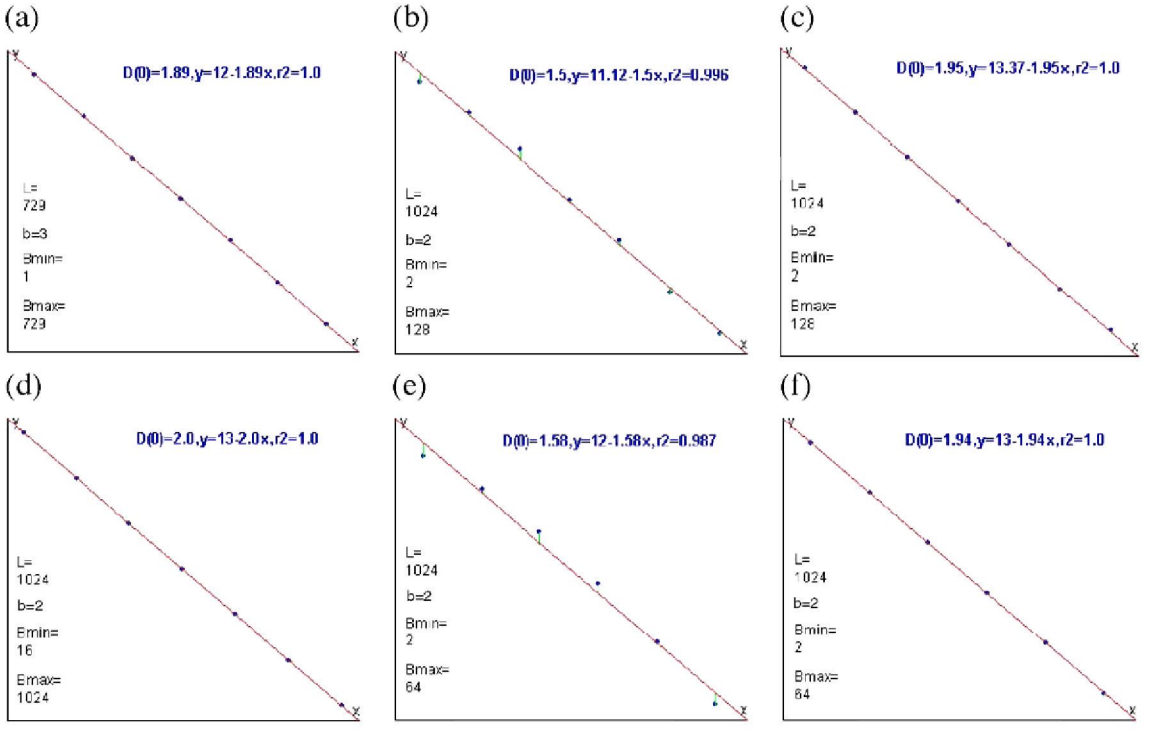

Fig. 3. Estimations of a fractal dimension $D(0)$ using the computer algorithm, to compare to the theoretical ones or to the ones already measured on real data. a) Sierpinski: The theoretical value is $\log 8/\log 3$. b) PSF4Interface: The theoretical value is $D(0) = \log 8/\log 4 = 1.5$ (Perrier et al., 1999). c) PSF4SolidMass: The theoretical mass "fractal" dimension is $D(0) = 2$ (Perrier et al., 1999). d) MFp_1p_2: The support is Euclidean ($D = 2$). The theoretical value is $D(0) = 2$. e) SEM3Interface: The value estimated by Dathe and Thullner (2005) is $D(0) = 1.53$. f) SEM3SolidMass: The value estimated by Dathe and Thullner (2005) is $D(0) = 1.94$.

The theoretical value of the MFp_1p_2 image, as well as the estimated one, is $D(0)=2$ for this Euclidean, non-fractal object (Fig. 3d).

Results for SEM3Interface and SEM3SolidMass images (Fig. 3e and f) are compared with results obtained by Dathe and Thullner (2005) with a different box-counting code. The estimated value of the mass fractal dimension is $D(0)=1.94$ exactly the same as in the Dathe and Thullner (2005) work. The discrepancy observed for the interface fractal dimension ($D(0)=1.58$ instead of 1.53) is explained by differences in boundary extraction algorithms. Dathe and Thullner (2005) extract the set of boundaries in the interface image as lines of 2 pixel width (one pixel in the solid space, the other one in the pore space), whereas we work here on the 1-pixel lines depicting the pore–solid boundaries.

Fig. 4 shows the spectrum of dimensions $D(q)$, $-10<q<10$ for each image. The first three images (Sierpinski, PSF4Interface, PSF4SolidMass) exhibit flat spectra, as expected. We used box sizes covering the whole size range of the images, and successive sizes according to the bases (3 and 4) respectively used in the construction of these theoretical models. But the value of $D(q)$ is smaller than the expected value $D=2$ for the PSF4SolidMass. The second set of three images (MFp_1p_2, SEM3Interface, SEM3SolidMass) exhibits decreasing spectra as expected for multifractal sets. For the synthetic image MFp_1p_2, we can compare the numerical estimates to the theoretical ones calculated using Eqs. (5a) and (5b), for $-10<q<10$. One can check in Fig. 4d that the theoretical and calculated spectra almost coincide.

4. Discussion and conclusions

When we work with images from the real world instead of mathematical models, we cannot rely on any theoretically based calculation of the fractal, non-fractal or multifractal behaviour. The difficulties associated with real images are that they have a finite size, their scale boundaries are defined by the image resolution, and we use numerical log/log fits (which tend to straighten any curve) to obtain the Rényi dimensions.

Generalized dimensions $D(q)$ can be obtained with the method of moments for any image and box size

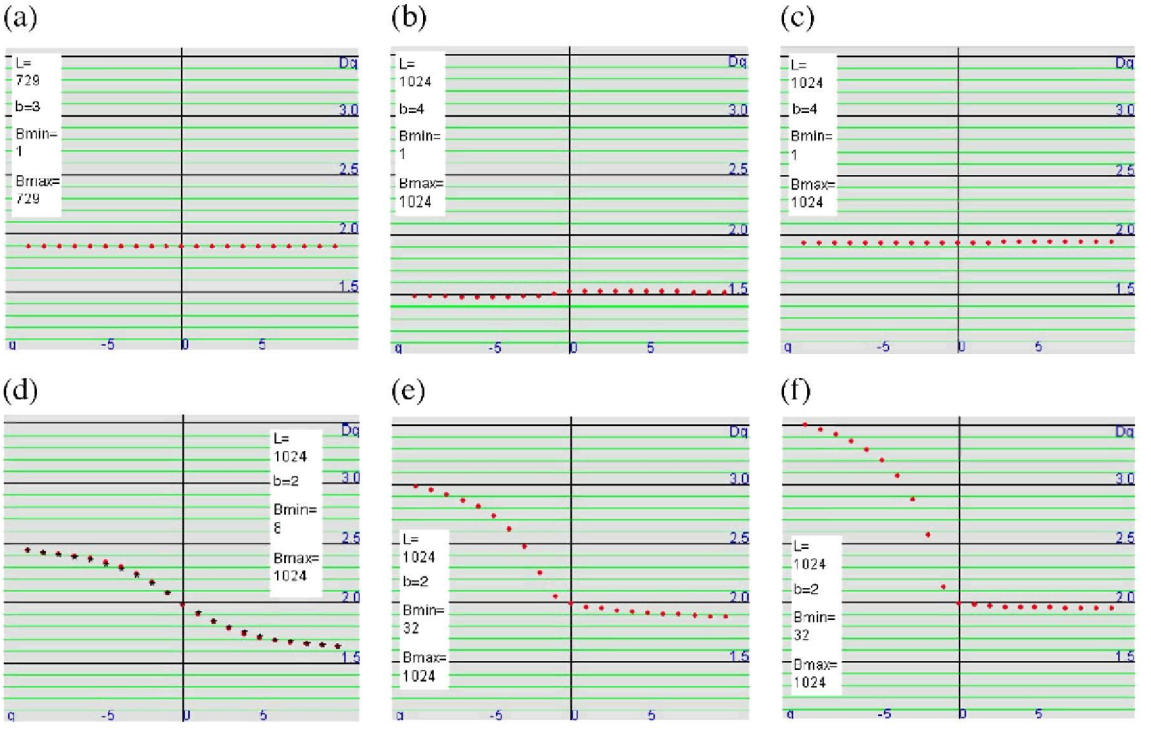

Fig. 4. Estimation of the spectra of generalized fractal dimensions for $-10<q<10$. The $D(q)$ in the calculated spectra (red dots) have been obtained with regression fits where the coefficient of determination R^2 was always higher than 0.82. a) Sierpinski Carpet: The theoretical spectrum is line $y=\log 8/\log 3=1.89$. $R^2>0.999$ for each q. b) PSF4Interface: The theoretical spectrum is line $y=1.5$. $R^2>0.99$ for $q>0$ and $R^2>0.82$ for $q<0$. c) PSF4SolidMass: The theoretical spectrum is line $y=2$. $R^2>0.999$ for each q. d) MFp_1p_2: The theoretical spectrum is plotted using star symbols. $R^2>0.999$ for $q>0$ and $R^2>0.998$ for $q<0$. e) SEM3Interface: $R^2>0.999$ for $q>0$ and $R^2>0.987$ for $q<0$. f) SEM3SolidMass: $R^2>0.999$ for $q>0$ and $R^2>0.967$ for $q<0$. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

range. The software depicted in the present paper provides a tool for comparing the resulting $D(q)$ for different box size ranges. It enables a user to avoid inappropriate ranges that show an increasing spectrum of generalized dimensions. The software may also help to analyze spectra obtained for different ranges of box sizes and different bases for the successive box sizes used to cover the measured object in the algorithm.

Even with synthetic images built from simple mathematical models, the range of box sizes used in the numerical analysis is a key issue. On the PSF images as well as on the MF image, we checked the numerical values calculated at each algorithm step. When similar conditions (base, box size range) are used, the theoretical and numerical data sets are identical, but only on the same finite range of scale. Outside of this scale range, the theoretical values of $D(q)$ (calculated as the limit when r approaches zero) and the numerical values of $D(q)$ (calculated from the regression fits) may of course differ. They are still very close if we select the range of scale and a grid matching the theoretical generation pattern. They may differ greatly if we use an arbitrary base and box size range value. This is illustrated in Fig. 5.

For the theoretical example MFp_1p_2, using base 2 (Fig. 5a), a maximum box size equal to 1024 pixels, and decreasing minimum box sizes from 64 to 1, the calculated $D(q)$ remain close to the theoretical ones, except when the minimum box size goes below 8. For the range 1–256, the $D(q)$ are even increasing with respect to q, and the coefficient of determination is the lowest ($R^2=0.86$), which indicate that the latter trials become unreliable. With base 3 (Fig. 5b), the fitted $D(q)$ exhibit a trend comparable to the mathematical ones, but nevertheless need to be rejected when compared to the best estimation shown previously in Fig. 4c. The unsatisfactory results obtained for the MFp_1p_2 image using small box sizes in the multifractal algorithm can be explained by the construction of this image, which has been obtained after only 8 iterations, as described in Stanley and Meakin (1988). Inside the smallest boxes, the variability of the mass is no more governed by a binomial distribution but nearly a random (Poisson) distribution. We could build synthetic images using a larger number of iterations to better approach the mathematical model, but we would like to point out that a similar uncertainty in the mass spatial variability is likely to occur at the smallest observable scales measured in real images.

The scatter of the calculated spectra is very low for the theoretical example PSF4SolidMass, using base 2 (Fig. 5c), a maximum box size equal to 1024 pixels, and decreasing minimum box sizes from 64 to 1. The coefficient of determination is always higher than 0.99. With base 3 (Fig. 5d), the larger the range of the box size, the higher the significance of the coefficient of determination, and the closer the spectrum to the expected flat one with values close to 2. So even using a grid which does not match the underlying pattern, we can approach acceptable numerical results.

To conclude, this paper is a contribution in the wide scientific research field of the comparison of discrete computer algorithms versus mathematical theory. When the computer software will process real images, the user is strongly encouraged to carry out all possible sensitivity analyses and statistical tests to evaluate the box size range effect and to assess the quality of the log/log fits involved in the numerical multifractal analysis.

Acknowledgments

The funding for one of the authors (AD) from the German Research Foundation under grant number DA 575/1-2, and more recently from Cornell's Biogeochemistry and Biocomplexity Initiative is greatly appreciated. The authors like to thank a lot Chantal Bernard from IRD, who prepared the Webpage http:///www.bondy.ird.fr/~perrier/multifractals/MF.htm, and Cedric Aubouy who helped to have the applet working distantly through the Internet.

Appendix 1. Theoretical calculation of the multifractal spectrum for the solid mass in the PSF model

Let us consider a representative volume of porous medium of linear size L and fractal dimension D, in a space of Euclidean dimension d. The total solid mass of a sample of size L^d and porosity Φ is:

$$M = (1-\Phi)L^d \tag{A1.1}$$

Let us recall that the PSF model is made of S, P and F respectively for Solid, Pore and Fractal porous subparts. When we look at the mass of a given S unit of linear size r, its mass is

$$M_{is}(r) = r^d \tag{A1.2}$$

and, since $p_i(r)=M_i(r)/M$ from Eq. (1), one gets:

$$p_{is}(r) = \frac{1}{(1-\Phi)}\left(\frac{r}{L}\right)^d \tag{A1.3}$$

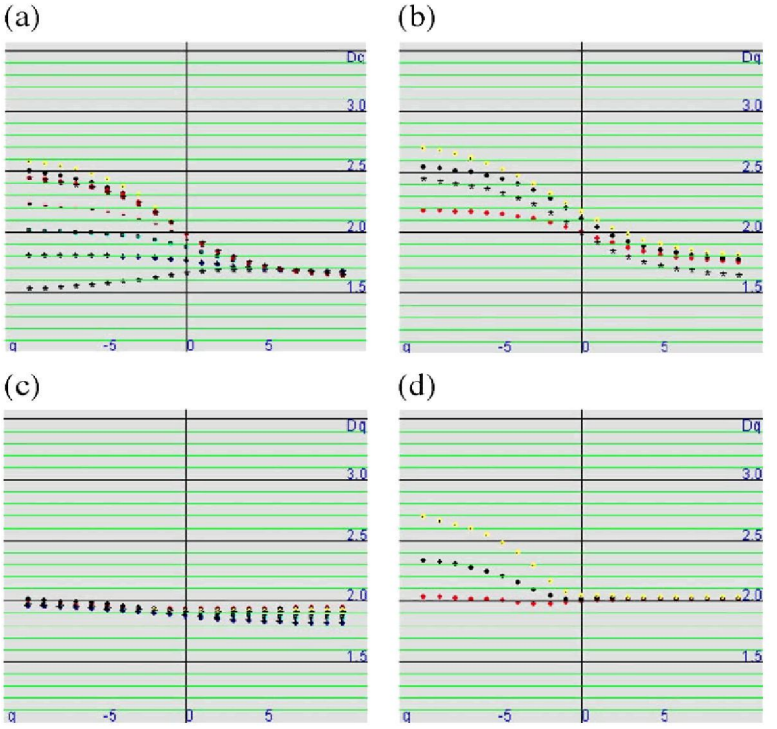

Fig. 5. Several estimated multifractal spectra associated to a single theoretical multifractal spectrum. a) MFp_1p_2, base 2: light red (64-1024), black (32-1024), yellow (16-1024), red (8-1024), pink (4-1024), light blue (2-1024), dark blue (1-1024), grey (1-256). b) MFp_1p_2, base 3: red (3-729), black (9-729), yellow (27-729), * (theoretical). c) PSF4SolidMass, base 2: red (1-1024), black(2-1024), yellow (4-1024), grey (8-1024), pink (16-1024), light blue (32-1024), dark blue (64-1024). d) PSF4SolidMass, base 3: from the top curve to the bottom one, yellow (9-729), black(3-729), red (1-729). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

From Perrier et al. (1999) we know that the porosity of each subpart F of any linear size r is identical to the total porosity, due to fractal self-similarity, so:

$$p_{if}(r) = \left(\frac{r}{L}\right)^d$$

From Eq. (9) in Perrier et al. (1999), we know that the total number of F subparts is:

$$N_f(r) = L^D r^{-D} \tag{A1.4}$$

and from Eq. (44a) in Perrier et al. (1999), we know that the number $B_S(r)$ of boxes of linear size r needed to cover the solid whole space in the PSF model is:

$$B_S(r) = (1-\Phi)L^d r^{-d} + \Phi L^D r^{-D} \tag{A1.5}$$

It follows from the Eq. (A1.5), that we need at each scale $B_S(r)=(1-\Phi)L^d r^{-d}+\Phi L^D r^{-D}$ boxes to cover the whole solid space, this includes $N_f=L^D r^{-D}$ boxes (Eq. (A1.4)) needed to cover the porous fractal part F, so $N_S=(1-\Phi)(L^d r^{-d}-L^D r^{-D})$ boxes are needed to cover the "full" solid parts S of mass r^d.

$$N_S = (1-\Phi)(L^d r^{-d}-L^D r^{-D}) \tag{A1.6}$$

Let us now calculate the Renyi dimensions:

For $q \neq 1$, from Eq. (2), $D(q) = \lim_{r\to 0} \frac{1}{q-1} \frac{\log \sum_{i=1}^{N(r)} p_i^q(r)}{\log r}$

$$\sum_{i=1}^{N(r)} p_i^q(r) = \sum_{i=1}^{N_S(r)} p_{is}^q(r) + \sum_{i=1}^{N_f(r)} p_{if}^q(r) = N_S\, p_{is}^q(r) + N_f\, p_{if}^q(r) = (1-\Phi)(L^d r^{-d} - L^D r^{-D})\left(\frac{1}{(1-\Phi)}\left(\frac{r}{L}\right)^d\right)^q$$

$$+L^D r^{-D}\left(\left(\frac{r}{L}\right)^d\right)^q = \left(\frac{r}{L}\right)^{d(q-1)}\left((1-\Phi)^{1-q}\left(1-\left(\frac{r}{L}\right)^{d-D}\right)+\left(\frac{r}{L}\right)^{d-D}\right)$$

$$D_S(q) = \frac{1}{q-1}\lim_{r\to 0}\frac{d(q-1)\log r - d(q-1)\log L + \log\left((1-\Phi)^{1-q}\left(1-\left(\frac{r}{L}\right)^{d-D}\right)+\left(\frac{r}{L}\right)^{d-D}\right)}{\log r}$$

$$D_S(q) = d + \lim_{r\to 0}\frac{-d(q-1)\log L + \log\left((1-\Phi)^{1-q}\left(1-\left(\frac{r}{L}\right)^{d-D}\right)+\left(\frac{r}{L}\right)^{d-D}\right)}{(q-1)\log r}$$

Since $\lim_{r\to 0}\left(\frac{r}{L}\right)^{d-D} = 0$ and $\lim_{r\to 0}\log r = \infty$ thus

$$D_S(q) = d \tag{A1.7}$$

For $q=1$, from Eq. (3), $D(1) = \lim_{r\to 0}\dfrac{\sum_{i=1}^{N(r)} p_i(r)\log p_i(r)}{\log r}$

$$\sum_{i=1}^{N(r)} p_i(r)\log p_i(r) = \sum_{i=1}^{N_S(r)} p_{is}(r)\log p_{is}(r) + \sum_{i=1}^{N_f(r)} p_{if}(r)\log p_{if}(r) = N_S p_{is}(r)\log p_{is}(r) + N_f p_{if}(r)\log p_{if}(r)$$

$$= (1-\Phi)(L^d r^{-d} - L^D r^{-D})\frac{1}{(1-\Phi)}\left(\frac{r}{L}\right)^d \log\left(\frac{1}{(1-\Phi)}\left(\frac{r}{L}\right)^d\right) + L^D r^{-D}\left(\frac{r}{L}\right)^d \log\left(\frac{r}{d}\right)^d$$

$$= \left(1-\left(\frac{r}{L}\right)^{d-D}\right)\left(\log\left(\frac{r}{L}\right)^d - \log(1-\Phi)\right) + \left(\frac{r}{L}\right)^{d-D}\log\left(\frac{r}{L}\right)^d = \log\left(\frac{r}{L}\right)^d + \left(\left(\frac{r}{L}\right)^d - 1\right)\log(1-\Phi)$$

$$D_S(1) = \lim_{r\to 0}\frac{\log\left(\frac{r}{L}\right)^d + \left(\left(\frac{r}{L}\right)^d - 1\right)\log(1-\Phi)}{\log r} = \lim_{r\to 0}\left(d + \frac{-d\log L + \left(\left(\frac{r}{L}\right)^d - 1\right)\log(1-\Phi)}{\log r}\right)$$

As previously, one concludes that

$$D_S(1) = d \tag{A1.8}$$

One could show similarly that symmetrical results hold for the pore space:

$$\forall q,\ D_P(q) = D_S(q) = d \tag{A1.9}$$

Appendix 2. Theoretical calculation of the multifractal spectrum for the set MFp1p2 generated by a two-dimensional random multiplicative process

Let us consider a representative volume of linear size $L=1$, divided at each iteration k in 2^{2k} boxes of linear size $r=2^{-k}$

For each box of size r, the probability to be filled is $p_i(k) = \prod_{j=1}^{k} p_j$ where $p_j=p_1$ or $p_j=p_2$. The result of this multiplicative process is a binomial distribution of the mass in the 2^{2k} boxes of linear size r, where, for each j between 0 and k, $2^k\binom{j}{k}$ boxes of linear size $r=2^{-k}$ and area $r=2^{-2k}$ have the same probability $p_1^j p_2^{k-j}$ to be filled.

The total mass at iteration k is

$$M = \left(\sum_{j=1}^{j=k} 2^k\binom{j}{k} p_1^j p_2^{k-j}\right)2^{-2k} = 2^{-k}\sum_{j=1}^{j=k}\binom{j}{k} p_1^j p_2^{k-j} = 2^{-k}(p_1+p_2)^k \tag{A2.1}$$

And

$$p_i(r) = p_1^j p_2^{k-j}\frac{2^{-2k}}{M} = \frac{p_1^j p_2^{k-j}}{2^k(p_1+p_2)^k} \tag{A2.2}$$

for each of the k sets made of similar $2^k\binom{j}{k}$ boxes.

For $q\neq 1$, from Eq. (2), $D(q)=\lim_{r\to 0}\frac{1}{q-1}\frac{\log\sum_{i=1}^{N(r)}p_i^q(r)}{\log r}$, so:

$$\sum_{i=1}^{N(r)}p_i^q(r)=2^k\sum_{j=1}^{j=k}\binom{j}{k}\left(\frac{p_1^j p_2^{k-j}}{2^k(p_1+p_2)^k}\right)^q=\frac{2^k}{(2^k(p_1+p_2)^k)^q}\sum_{j=1}^{j=k}\binom{j}{k}(p_1^q)^j(p_2^q)^{k-j}=\frac{1}{2^{k(q-1)}(p_1+p_2)^k}(p_1^q+p_2^q)^k$$

$$D(q)=\lim_{r\to 0}\frac{1}{q-1}\log\frac{\sum_{i=1}^{N(r)}p_i^q(r)}{\log r}=\lim_{k\to\infty}\frac{1}{q-1}\frac{\log\frac{(p_1^q+p_2^q)^k}{2^{k(q-1)}(p_1+p_2)^k}}{\log 2^{-k}}=\lim_{k\to\infty}\frac{1}{q-1}\frac{k\log\frac{(p_1^q+p_2^q)}{2^{(q-1)}(p_1+p_2)}}{-k\log 2}=\frac{\log\left(\frac{p_1^q+p_2^q}{2^{q-1}(p_1+p_2)^q}\right)}{(1-q)\log(2)} \tag{A2.3}$$

$$D(q)=1+\frac{\log(p_1^q+p_2^q)-q\log(p_1+p_2)}{(1-q)\log(2)}$$

This is the proof of Eq. (5a) in the text of the paper.

For $q=1$, from Eq. (3), $D(1)=\lim_{r\to 0}\frac{\sum_{i=1}^{N(r)}p_i(r)\log p_i(r)}{\log r}$, so:

$$D(1)=\lim_{k\to\infty}\frac{2^k\sum_{j=1}^{j=k}\binom{j}{k}\left(\frac{p_1^j p_2^{k-j}}{2^k(p_1+p_2)^k}\right)\log\left(\frac{p_1^j p_2^{k-j}}{2^k(p_1+p_2)^k}\right)}{\log(2^{-k})}$$

$$D(1)=\lim_{k\to\infty}\frac{1}{-k(p_1+p_2)^k\log(2)}\sum_{j=1}^{j=k}\binom{j}{k}(p_1^j p_2^{k-j}\log(p_1^j p_2^{k-j})-p_1^j p_2^{k-j}\log(2(p_1+p_2))^k)$$

$$D(1)=\lim_{k\to\infty}\frac{\sum_{j=1}^{j=k}\binom{j}{k}\left(p_1^j p_2^{k-j}\log(p_1^j p_2^{k-j})\right)}{-k(p_1+p_2)^k\log(2)}+\frac{\log(2(p_1+p_2))^k}{k(p_1+p_2)^k\log(2)}\sum_{j=1}^{j=k}\binom{j}{k}(p_1^j p_2^{k-j})$$

$$D(1)=\lim_{k\to\infty}\frac{\log p_1\sum_{j=1}^{j=k}j\binom{j}{k}p_1^j p_2^{k-j}+\log p_2\sum_{j=1}^{j=k}(k-j)\binom{j}{k}p_1^j p_2^{k-j}}{-k(p_1+p_2)^k\log(2)}+\frac{\log(2(p_1+p_2))^k}{k(p_1+p_2)^k\log(2)}(p_1+p_2)^k$$

$$D(1)=\lim_{k\to\infty}\frac{kp_1\log p_1\sum_{j=1}^{j=k}\binom{j-1}{k-1}p_1^{j-1}p_2^{k-j}+kp_2\log p_2\sum_{j=1}^{j=k}\binom{j}{k-1}p_1^j p_2^{k-j-1}}{-k(p_1+p_2)^k\log(2)}+\frac{\log(2(p_1+p_2))^k}{k\log(2)} \tag{A2.4}$$

$$D(1)=\lim_{k\to\infty}\frac{p_1\log p_1\sum_{j=1}^{j=k-1}\binom{j}{k-1}p_1^j p_2^{k-1-j}+p_2\log p_2\sum_{j=1}^{j=k-1}\binom{j}{k-1}p_1^j p_2^{k-1-j}}{-(p_1+p_2)^k\log(2)}+\frac{\log(2(p_1+p_2))^k}{k\log(2)}$$

$$D(1)=\lim_{k\to\infty}\frac{(p_1\log p_1+p_2\log p_2)(p_1+p_2)^k}{-(p_1+p_2)^k\log(2)}+\frac{\log(2(p_1+p_2))^k}{k\log(2)}$$

$$D(1)=\frac{(p_1\log p_1+p_2\log p_2)}{-\log(2)}+\frac{\log(2(p_1+p_2))}{\log(2)}$$

$$D(1)=\frac{\log(2(p_1+p_2))-(p_1\log p_1+p_2\log p_2)}{\log(2)}$$

This is the proof of Eq. (5b) in the text of the paper.

References

Dathe, A., Thullner, M., 2005. The relationship between fractal properties of solid matrix and pore space in porous media. Geoderma 129, 279–290.

Dathe, A., Perrier, E., Tarquis, A., 2006-this issue. Multifractal analysis of two-dimensional soil porous structure on natural images. Geoderma 134, 318–326. doi:10.1016/j.geoderma.2006.03.024.

Stanley, H., Meakin, P., 1988. Multifractal phenomena in physics and chemistry. Nature 335, 405–409.

Perrier, E., Bird, N., Rieu, M., 1999. Generalizing the fractal model of soil structure: the pore-solid fractal approach. Geoderma 88, 137–164.